

Gene expression becomes heterogeneous with age

Mehmet Somel^{1*}, Philipp Khaitovich¹, Sabine Bahn², Svante Pääbo¹, and Michael Lachmann¹

Evidence from a number of fields indicates that phenotypic variation among individuals and among somatic cells increases with age [1–5]. However, no evidence for an age-dependent increase in variation in gene expression has yet been found [6–9]. Using eight microarray data sets from different studies in humans and rats, we find that gene expression becomes more variable with advancing age. This age-correlated heterogeneity of expression (ACHE) has a minor effect on individual genes, but is widespread throughout the transcriptome. We find no evidence for ACHE being caused by germline mutations or by differences in aging rates between individuals. However, our results are compatible with ACHE being an outcome of the accumulation of stochastic effects at the cellular level.

We designed a statistical test that measures an increase in expression heterogeneity with age. ACHE is calculated separately for each probe set (a probe set detects one mRNA transcript in microarray experiments). We tested eight microarray data sets, including data from human kidney (cortex and medulla), human skeletal muscle, post-mortem human cerebral cortex and rat hippocampus, for an excess of probe sets showing ACHE (Figure 1). The random expectation was calculated through permutations (see Supplemental Data published with this article online).

We found that the proportion of probe sets showing a positive correlation between age and

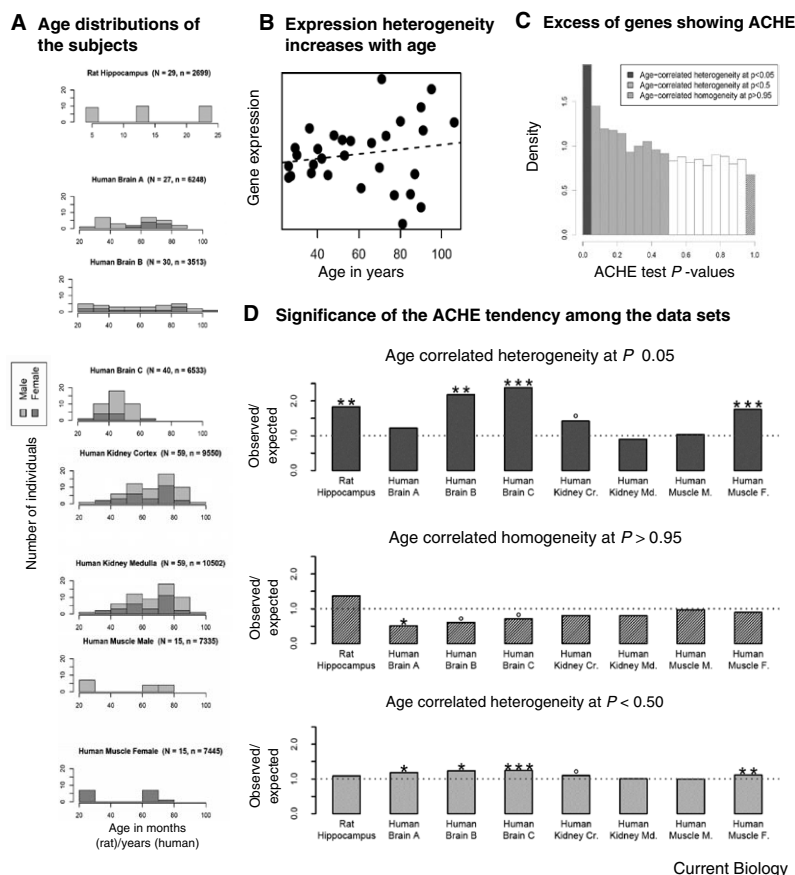


Figure 1. Data sets and results of the Age-Related Heterogeneity of Expression (ACHE) test.

(A) Age distribution of subjects in the data sets. Human ages are given in years, rat ages in months. In parentheses, the number of subjects is given as 'N', the number of probe sets used in quantitative ACHE analyses is given as 'n'. The ACHE test for increase in heterogeneity with age is conducted on each probe set independently and makes use of regression models between expression level and age [14] (Supplemental Data). (B) An example from the human brain data set B, the log-expression versus age plot for a probe set detecting the gene PIM-1, for which the ACHE test *p*-value was calculated as 0.0002. (C) ACHE test *p*-value distribution for human brain B. (D) shows the main result of the study for quantitative gene expression. The heights of the bars indicate observed to expected ratios of the number of probe sets at different cutoffs within the ACHE test *p*-value distribution. The observed number is the number of probe sets found within the regions of the *p*-value distributions at cutoffs $p \leq 0.05$, $p \geq 0.95$ and $p < 0.5$, from top to bottom. The expected number is the median number of probe sets found within 1000 random *p*-value distributions at the same cutoff, which represents the expectation under the null hypothesis. The dashed line shows observed to expected ratio = 1. Significance of the observed to expected ratios are indicated above the bars as (o) for $0.05 < p \leq 0.10$, as (*) for $0.01 < p \leq 0.05$, as (**) for $0.001 < p \leq 0.01$ and as (***) for $p \leq 0.001$. None of the results are significant in the direction of increase in homogeneity with age.

expression variation (at ACHE test $p \leq 0.05$) was more than expected in seven of the eight data sets (Figure 1D). In four of these data sets, the excess was significant (permutation test $p \leq 0.01$), resulting in an overall significance of $p = 7 \times 10^{-7}$ (Binomial test) across all eight data sets. Similarly, the proportion of

probe sets showing an increase in expression homogeneity at $p \geq 0.95$ was less than expected in seven of the eight experiments.

Additionally, we tested for heterogeneity in gene expression at the qualitative level (i.e. whether a gene is expressed or not). The proportion of probe sets showing a positive correlation between

age and qualitative expression variation at $p \leq 0.05$ was more than expected in seven of the eight data sets and three of these were significant (permutation test $p \leq 0.05$), resulting in an overall significance of $p = 0.006$ (Binomial test). The results from both the quantitative and the qualitative dimensions of gene expression indicate that ACHE is a weak but pervasive characteristic of the human and rat transcriptomes.

What might be causing ACHE? We first asked whether genes showing strong ACHE at $p \leq 0.05$ differ in a consistent way from other genes. We found that, first, there was no significant overlap between genes showing strong ACHE in different experiments. Second, there was no significant enrichment of strong ACHE genes in functional gene groups. Lastly, we observed a correlation between decrease in expression levels and increase in variation. However, ACHE was not restricted to genes undergoing age-related-decrease in expression levels (Supplemental Data). Taken together, these results imply that genes showing strong ACHE do not belong to a specific set of genes.

Medawar's germline mutation accumulation theory of aging [3] provides a possible explanation for ACHE. Because old individuals contribute less to the next generation than the young, the force of purifying selection on the old phenotype is weaker. Therefore, a deleterious germline mutation that affects gene expression only at old age will have a higher probability of remaining within the population than other mutations. What we are observing as ACHE could therefore be a signature of this old-age-specific genetic heterogeneity. However, a genetically homogenous rat population also displayed significant ACHE (Figure 1D). Thus, the germline mutation accumulation theory cannot fully account for the ACHE effect.

Another possible source of ACHE are differences in aging rates between individuals [10]. If expression profiles are similar

among young individuals, but then change at different rates due to genetic or environmental variation, the expression profiles of the old will be relatively heterogeneous. To evaluate this hypothesis, we assigned an 'expression age' to each individual for each gene, according to expression level [11] (Supplemental Data). Expression ages derived from independent sets of genes were tightly correlated, implying that, within an individual, the vast majority of genes change their expression levels with age at similar rates. However, we did not detect a significant age-related increase in the absolute differences between mean expression ages and chronological ages, which we would expect if differences in aging rates were reflected in gene expression.

Our results indicate that ACHE is a general — but weak — effect in the transcriptome. This is compatible with ACHE being the outcome of accumulating stochastic effects in the soma, such as cellular damage and mutations. These effects will influence each cell in a unique way [2] so that expression variation among aging cells will be equalized at the tissue level. If somatic mutations mostly cause decreases in expression level [12], the overlap between ACHE and age-related decrease in expression levels can also be explained within this framework. ACHE supports the stochastic nature attributed to the aging process [10]. It implies a weakening of expression regulation with age, contrary to previous observations and hypotheses based on measurements on a small number of genes [6,13]. The extent of ACHE among single cells and the role of ACHE in aging remains to be determined.

Acknowledgments

We thank R. Sonu, S. Welle and E. Blacklock for providing data, I. Hellman and M. Leinweber for assistance in analyses, and W. Enard, S. Myles, S. Ptak, K. Tang, C. Green and two anonymous reviewers for helpful suggestions.

Supplemental data

Supplemental data including Experimental Procedures and Acknowledgements are available at <http://www.current-biology.com/cgi/content/full/16/10/R359/DC1/>.

References

1. Nelson, E.A., and Dannefer, D. (1992). Aged heterogeneity: fact or fiction? The fate of diversity in gerontological research. *The Gerontologist* 32, 17–23.
2. Herndon, L.A., Schmeissner, P.J., Dudaronek, J.M., Brown, P.A., Listner, K.M., Sakano, Y., Paupard, M.C., Hall, D.H., and Driscoll, M. (2002). Stochastic and genetic factors influence tissue-specific decline in ageing *C. elegans*. *Nature* 419, 808.
3. Hughes, K.A., Alipaz, J.A., Drnevich, J.M., and Reynolds, R.M. (2002). A test of evolutionary theories of aging. *Proc. Natl. Acad. Sci. USA* 99, 14286–14291.
4. Fraga, M.F., Ballestar, E., Paz, M.F., Ropero, S., Setien, F., Ballestar, M.L., Heine-Suner, D., Cigudosa, J.C., Urioste, M., Benitez, J., et al. (2005). From The Cover: Epigenetic differences arise during the lifetime of monozygotic twins. *Proc. Natl. Acad. Sci. USA* 102, 10604–10609.
5. Rea, S.L., Wu, D., Cypser, J.R., Vaupel, J.W., and Johnson, T.E. (2005). A stress-sensitive reporter predicts longevity in isogenic populations of *Caenorhabditis elegans*. *Nat. Genet.* 37, 894.
6. Rogina, B., Vaupel, J.W., Partridge, L., and Helfand, S.L. (1998). Regulation of gene expression is preserved in aging *Drosophila melanogaster*. *Curr. Biol.* 8, 475.
7. Pletcher, S.D., Macdonald, S.J., Marguerie, R., Certa, U., Stearns, S.C., Goldstein, D.B., and Partridge, L. (2002). Genome-wide transcript profiles in aging and calorically restricted *Drosophila melanogaster*. *Curr. Biol.* 12, 712–723.
8. Welle, S., Brooks, A.J., Delehanty, J.M., Needler, N., and Thornton, C.A. (2003). Gene expression profile of aging in human muscle. *Physiol. Genomics* 14, 149–159.
9. Golden, T.R., and Melov, S. (2004). Microarray analysis of gene expression with age in individual nematodes. *Aging Cell* 3, 111–124.
10. Kirkwood, T.B.L. (2005). Understanding the Odd Science of Aging. *Cell* 120, 437.
11. Rodwell, G.E.J., Sonu, R., Zahn, J.M., Lund, J., Wilhelmy, J., Wang, L., Xiao, W., Mindrinos, M., Crane, E., Segal, E., et al. (2004). A transcriptional profile of aging in the human kidney. *PLoS Biol.* 2, e427.
12. Lu, T., Pan, Y., Kao, S.-Y., Li, C., Kohane, I., Chan, J., and Yankner, B.A. (2004). Gene regulation and DNA damage in the ageing human brain. *Nature* 429, 883.
13. Helfand, S.L., and Inouye, S.K. (2002). Rejuvenating views of the ageing process. *Nat. Rev. Genet.* 3, 149.
14. Faraway, J. (2002). Practical Regression and Anova Using R. <http://www.stat.lsa.umich.edu/~faraway/book>.

¹Department of Evolutionary Genetics, Max Planck Institute for Evolutionary Anthropology, Deutscher Platz 6, D-04103 Leipzig, Germany. ²Institute of Biotechnology, University of Cambridge, Cambridge CB2 1QT, UK.

*E-mail: somel@eva.mpg.de